

model produces a score, which is a measure of the likelihood of premium fraud or abuse. The predictive model is included in a system that accepts policies to be considered for scoring, selects which policies are appropriate for scoring, stores data about the policies in a database, uses the data to derive variables for the model, and processes and outputs the model scores and related information. A rule-based analysis, which detects specific inconsistencies in the data that are indicative of premium fraud, may also be part of the system. The model scores and red-flag indicators from the rule-based analysis may be further processed to provide customized output for users.

#### IN THE CLAIMS:

Please cancel claims 42-44. Please amend claims 2, 5-8, 14, and 58-61. For the Examiner's convenience, all claims are reproduced below. A marked-up version of the amendments made herein follows the Remarks section.

1           1.       A method for detecting misrepresentation of policy related information  
2 provided to an insurer by a policyholder where the information is used by the insurer in  
3 determining an amount of premium to be paid for insurance coverage provided to the  
4 policyholder, the method comprising:  
5           selecting a plurality of insurance policies to process with a predictive model;  
6           for each selected policy, deriving variables from policy related information provided by  
7           the policyholder in connection with the selected policy; and  
8           for each selected policy, applying the derived variables of the policy to the predictive  
9           model to generate a model score indicating the relative likelihood of misrepresented  
10          information provided by the policyholder or an expected adjustment of the  
11          premium on the policy.

1        2. (Amended) The method of claim 1, further comprising:  
2        collecting training data including a plurality of insurance policies having  
3        misrepresented information and a plurality of policies not having misrepresented  
4        information;  
5        developing the predictive model from the training data; and  
6        storing the predictive model.

1        3.        The method of claim 1, further comprising:  
2        converting the model score to a fraud score indicating a probability of fraud in the  
3        policy.

1        4.        The method of claim 1, further comprising:  
2        converting the model score to the expected adjustment of the premium on the policy.

1        5. (Amended) The method of claim 1, wherein selecting a plurality of insurance  
2        policies further comprises:  
3        for each policy, automatically determining start and end dates of a scoring period over  
4        which misrepresented policy information is to be detected.

1        6. (Amended) The method of claim 5, wherein the start and end dates of the scoring  
2        period include a period for which the policy has consistent and complete data.

1        7. (Amended) The method of claim 5, further comprising:  
2        responsive to a policy not having consistent or complete data in the scoring period,  
3        defining an exclusion code providing a reason that the policy was not selected.

1           8. (Amended) The method of claim 5, wherein the insurance policies are workers'  
2   compensation insurance policies, and automatically determining start and end dates of the  
3   scoring period further comprises:

4           defining the start and end dates such that all audit adjustments are contained between  
5           the start and end dates.

1           9.     The method of claim 1, wherein selecting a plurality of insurance policies further  
2   comprises:

3           for each policy, receiving a user defined scoring period to be scored for the policy; and  
4           automatically selecting those policies having consistent and complete data in the  
5           respective user defined time period from which the variables for the predictive  
6           model may be derived.

1           10.    The method of claim 9, further comprising:  
2           responsive to a policy not having consistent or complete data in the user defined time  
3           period defining an exclusion code providing a reason that the policy was not  
4           selected.

1           11.    The method of claim 9, further comprising:  
2           responsive to a policy not having consistent or complete data in the user defined scoring  
3           period, automatically suggesting a scoring period in which the policy has consistent  
4           and complete data.

1           12.    The method of claim 1, wherein deriving variables from policy related  
2   information further comprises:  
3           determining a plurality of peer groups of which the selected policy is a member; and

4 for each peer group or set of peer groups of which the selected policy is a member,  
5 deriving variables from the policy information which attribute characteristics of the  
6 peer group or set of peer groups to the selected policy, or which compare the  
7 selected policy to other policies in the peer group or set of peer groups.

1 13. The method of claim 12, wherein the derived variables estimate the probability  
2 of a dichotomous outcome or a certain distributional statistic of a continuous quantity for a  
3 policy, based on the peer group(s) of which the policy is a member.

1 14. (Amended) The method of claim 12, wherein deriving variables for the policy  
2 which compare the policy to other policies in its peer group(s) further comprises deriving  
3 variables that compare either at least one characteristic of the policy with at least one  
4 corresponding characteristic of the policies in its peer group(s).

1 15. The method of claim 12, further comprising:  
2 for each of the plurality of peer groups, storing in a lookup table group statistics for  
3 policy characteristics of the policies in the peer group; and  
4 deriving the variables for a selected policy by determining the peer group to which the  
5 selected policy belongs and using the statistics for the policy characteristics for the  
6 peer group to derive the variables for the selected policy.

1 16. The method of claim 15, further comprising:  
2 updating the lookup table for a peer group of the selected policy using policy  
3 information from the selected policy.

1        17.    The method of claim 1, wherein deriving variables further comprises:  
2        deriving variables from the policy information which compare the selected policy in a  
3        selected time period with the selected policy in a time period prior to the selected  
4        time period.

1        18.    The method of claim 17, wherein deriving variables from the policy information  
2        which compare the selected policy in a selected time period with the selected policy in a time  
3        period prior to the selected time period further comprises:  
4        deriving variables which quantify an amount or distribution of risk-related activities  
5        associated with the policy.

1        19.    The method of claim 17, wherein deriving variables from the policy information  
2        which compare the selected policy in a selected time period with the selected policy in a time  
3        period prior to the selected time period further comprises:  
4        determining at least one measure which is a percentage change in a policy characteristic  
5        between the selected time period and the previous time period.

1        20.    The method of claim 17, wherein deriving variables from the policy information  
2        which compare the selected policy in a selected time period with the selected policy in a time  
3        period prior to the selected time period further comprises:  
4        determining a vector of policy characteristics for the selected time period and a vector of  
5        the policy characteristics in the prior time period; and  
6        determining a scalar measure of comparison between the two vectors.

1        21.    The method of claim 20, wherein the scalar measure of comparison between the  
2        two vectors is computed as either a measure of distance between the two vectors or an angle  
3        measure between the two vectors.

1           22.     The method of claim 17, wherein deriving variables from the policy information  
2     which compare the selected policy in a selected time period with the selected policy in a time  
3     period prior to the selected time period further comprises:

4                 determining a percent change in a payroll share in at least one employment  
5                 classification in the selected time period relative to the previous time period.

1           23.     The method of claim 17, wherein deriving variables from the policy information  
2     which compare the selected policy in a selected time period with the selected policy in a time  
3     period prior to the selected time period further comprises:

4                 determining a percent change in a payroll share in an exception group in the selected  
5                 time period relative to the previous time period.

1           24.     The method of claim 17, wherein deriving variables from the policy information  
2     which compare the selected policy in a selected time period with the selected policy in a time  
3     period prior to the selected time period further comprises:

4                 determining a vector distance between vectors of payroll percent shares in each of a  
5                 plurality of employment classes in the selected time period and in the prior time  
6                 period.

1           25.     The method of claim 24, wherein the employment classes are SIC employment  
2     classes.

1           26.     The method of claim 24, wherein the employment class groups are NCCI  
2     employment class groups.

1           27.     The method of claim 24, wherein the employment class groups are rate-driven  
2     employment class groups.

1           28.     The method of claim 24, wherein the employment class groups are data-driven  
2     employment class groups, each group including employment classes that are likely to appear  
3     together in payroll reports.

1           29.     The method of claim 17, wherein deriving variables from the policy information  
2     which compare the selected policy in a selected time period with the selected policy in a time  
3     period prior to the selected time period further comprises:

4                 determining a percent change in a number of claims filed on the policy in the selected  
5                 time period relative to number of claims filed on the policy in the prior time period.

1           30.     The method of claim 17, wherein deriving variables from the policy information  
2     which compare the selected policy in a selected time period with the selected policy in a time  
3     period prior to the selected time period further comprises:

4                 determining a vector distance between a first vector of the number of claims filed in the  
5                 selected time period for each of a plurality of injury types and a second vector of the  
6                 number of claims filed in the prior time period in each of the plurality of injury  
7                 types.

1           31.     The method of claim 1, wherein the insurance policies are workers'  
2     compensation insurance policies and the policy relative information from which the variables  
3     for assessing the policies are derived includes payroll reports for the policyholder.

1           32.     The method of claim 1, further comprising:  
2     deriving direct policy variables which measure characteristics of the policyholder or the  
3     policy itself without comparison to other policies or the same policy in a prior time  
4     period.

1           33.     The method of claim 32 wherein the direct policy variables are selected from the  
2 group consisting of:

3           type of company of the policyholder;

4           location of the policyholder;

5           number of employees of the policyholder;

6           number of policy cancellations;

7           age of the policy;

8           industry type of the policyholder;

9           amount of payroll reported by the policyholder; and

10          distribution of payroll reported by the policyholder with respect to at least one

11          employment class.

1           34.     The method of claim 1, further comprising:

2           deriving direct claim variables which measure characteristics of claims filed on policy.

1           35.     The method of claim 34 wherein the direct claim variables are selected from the  
2 group consisting of:

3           number of claims filed during the selected time period;

4           dollar amount of claims filed during the selected time period;

5           type of claims filed during the selected time period;

6           number of claims filed during the selected time period relative to amount of premium

7           paid during the selected time period; and

8           number of claims filed during the selected time period relative to a size of payroll

9           during the selected time period.



1        36.     The method of claim 1, further comprising deriving variables that measure the  
2     probability of fraud in the policy conditionally based on at least one policy characteristic of the  
3     policy.

1        37.     The method of claim 1, further comprising:  
2        applying the policy to a plurality of decision rules which identify specific inconsistent or  
3        suspicious policy facts related to the policy, to generate an output indicating which  
4        decision rules were violated by the policy.

1        38.     The method of claim 37, wherein the decision rules are derived from statistical  
2     analysis of insurance policies of at least one insurer which have been determined to contain  
3     misrepresented policy information.

1        39.     The method of claim 37, wherein the insurance policies are workers'  
2     compensation insurance policies and wherein the decision rules are selected from a group  
3     consisting of:  
4        a decision rule that identifies as potentially fraudulent a policy that has an employment  
5        class code on a claim with an injury date during the selected time period but the  
6        employment class code for the claim is not included in payroll reports for the policy  
7        during the selected time period;  
8        a decision rule that identifies as potentially fraudulent a policy that reports zero payroll  
9        during the selected time period but for which one or more certificates of insurance  
10       were issued during the selected time period;  
11       a decision rule that identifies as potentially fraudulent a policy that reports zero payroll  
12       during the selected time period but which has at least one claim with an injury date  
13       during the selected time period;

14 a decision rule that identifies as potentially fraudulent a policy with an officer who is  
15 currently or was selectedly an officer on a different policy and where the new policy  
16 has a lower experience modification factor than the prior policy; and  
17 a decision rule that identifies as potentially fraudulent a policy that has an employment  
18 class code on a claim and for which no premium was reported at the time the claim  
19 was opened

1 40. The method of claim 1, further comprising:  
2 for each selected policy, determining at least one variable which significantly  
3 contributes to the model score for the policy; and  
4 outputting a reason for the model score with respect to the determined at least one  
5 variable.

1 41. The method of claim 40, wherein the insurance policies are workers'  
2 compensation insurance policies, and wherein the significant variable is selected from a group  
3 consisting of:

4 an indication of whether the policy has been previously audited;  
5 an indication of whether a reported payroll has been adjusted;  
6 a number of employment class codes in at least one payroll report of the policyholder  
7 during the selected time interval;  
8 a type of company of the policyholder;  
9 an age of the policy;  
10 a size of payroll of the policyholder;  
11 a size of a premium paid on the policy;  
12 an industry classification code of the policyholder;  
13 a distribution of payroll in at least one payroll report of the policyholder during the  
14 selected time interval;

15 a percent payroll share in a low rated employment class code;  
16 a change in a distribution of payroll in at least one payroll report of the policyholder  
17 during the selected time interval compared with the prior time period;  
18 a change in an exception group payroll share in at least one payroll report of the  
19 policyholder during the selected time interval compared with the prior time period;  
20 a payroll share in a group of agriculture related employment classes;  
21 a payroll share in a group of construction related employment classes;  
22 a payroll share in a group of manufacturing related employment classes;  
23 a payroll share in a group of government related employment classes;  
24 a payroll share in at least one clerical employment classes;  
25 a number of prior cancellations of the policy;  
26 a ratio of the number of claims made on the policy to a size of the payroll of the  
27 policyholder; and  
28 a number of claims on the policy during the selected time interval.

1 42. A method for training a neural network on a plurality of observations to score  
2 the observations on a dependent variable, each observation including an independent variable  
3 having an original value that is highly correlated with the dependent variable, so as to calibrate  
4 the influence of the independent variable on scores, the method comprising:

5 for each of the plurality of observations, setting the independent variable to a randomly  
6 selected value, and providing the observations to the neural network a first time,  
7 wherein the neural network establishes connection weights based on the provided  
8 observations to output an un-calibrated score for an observation; and  
9 for each of the plurality of observations, setting the independent variable to its original  
10 value in the observation, and providing the observations to the neural network a  
11 second time, wherein the neural network adjusts the connection weights to calibrate  
12 the output scores with respect to the independent variable.

1           43.     The method of claim 42, wherein the independent variable is a Boolean variable  
2     having two defined values, and the randomly set value is between the two defines values of the  
3     Boolean variable.

1           44.     The method of claim 42, wherein the independent variable is a continuous  
2     variable having a range of values, and the randomly set value is within the range of values.

1           45.     A method of estimating a quantity corresponding to a set of entities grouped  
2     using one or more hierarchical categories, the method comprising:  
3         determining an estimate of the quantity for a first category corresponding to the highest  
4         level of the hierarchy; and  
5         for each subsequent category representing a current, lower level of the hierarchy,  
6         adjusting the estimate of the quantity using an estimate for the current level and the  
7         estimate of the higher level.

1           46.     The method of claim 45, wherein the quantity being estimated is a risk factor,  
2     and each category of the hierarchy has a value for the risk factor.

1           47.     The method of claim 45, wherein the hierarchy of categories are Standard  
2     Industry Classification codes (SIC), and the quantity being estimated is risk factor associated  
3     with each SIC code.

1           48.     The method of claim 45, wherein adjusting the estimate of the quantity  
2     comprises applying a Bayesian adjustment to the estimate using the estimate for the current  
3     level of the hierarchy and the estimate of the quantity from the higher level.  
4

1        49.     A system for detecting premium fraud in an insurance policy, comprising:  
2        a database of insurance policies, each policy associated with a policyholder and having  
3        policy related data;  
4        a policy selection process that selects from the database a number of policies for scoring;  
5        a variable derivation process that derives for each of the selected policies variables  
6        associated with the policyholder of the policy for comparing the policy to peer  
7        group policies, and variables for comparing the policy in a selected time period with  
8        the policy a time period prior to the selected time period; and  
9        a fraud detection module that receives for each policy the derived variables and  
10       generates a score indicating the likelihood of misrepresentation of policy  
11       information by the policyholder of the policy.

1        50.     The method of claim 49, wherein the fraud detection module further comprises:  
2        a predictive model that generates a model score indicating a relative likelihood of  
3        misrepresentation of policy information by the policyholder; and  
4        a post scoring process that converts the model score into the fraud score indicating a  
5        probability of misrepresentation of policy information.

1        51.     The system of claim 50, wherein the post scoring process converts the model  
2        score into an expected adjustment of premium for a policy.

1        52.     The system of claim 50, further comprising:  
2        a rule-based process that applies a plurality of rules to a selected policy to identify  
3        policies suspected of premium fraud based on inconsistent or incomplete policy  
4        related information.

1        53.     A method for determining a usage strategy for processing insurance policies  
2     suspected of premium fraud, the suspected policies selected from a plurality of insurance  
3     policies, the method comprising:

4        establishing a frequency for scoring the plurality of insurance policies to obtain for each  
5        policy a score indicating a relative likelihood of premium fraud in the policy;  
6        establishing a ranking function for ranking the scored policies; and  
7        establishing a plurality of threshold scores, and for each threshold score, defining an  
8        audit action for performing on policies which have a score exceeding the threshold  
9        score, but not exceeding a next greater threshold score.

1        54.     The method of claim 53, wherein establishing a ranking function for ranking the  
2     scored policies further comprises:  
3        ranking the scored policies according to their scores.

1        55.     The method of claim 53, wherein establishing a ranking function for ranking the  
2     scored policies further comprises:  
3        ranking the scored policies according to an expected adjusted premium.

1        56.     The method of claim 53, wherein establishing a plurality of threshold scores  
2     further comprises:  
3        establishing a first threshold score for selecting for a desk audit those policies having a  
4        score exceeding the first threshold score; and  
5        establishing a second threshold score for selecting for a field audit those policies having  
6        a score exceeding the second threshold score, wherein the second threshold score is  
7        greater than the first threshold score.

1        57.     The method of claim 53, further comprising:  
2        establishing a set of rules for identifying policies suspected of premium fraud.

1        58. (Amended) The method of claim 53, further comprising:  
2        establishing a plurality of reason codes, each reason code providing an explanation for a  
3        policy receiving a score; and  
4        establishing for each of a number of reason codes, at least one audit action to be taken in  
5        response to a policy having a score which produces the reason code.

1        59. (Amended) A method for processing insurance policies suspected of premium  
2        fraud, the method comprising:  
3        scoring each of a plurality of insurance policies with a predictive model to generate for  
4        each policy a score indicating a relative likelihood of premium fraud;  
5        ranking the scored policies according to the scores;  
6        selecting for a desk audit those policies having a score exceeding a first threshold score;  
7        and  
8        selecting for a field audit those policies having a score exceeding a second threshold  
9        score, wherein the second threshold score is greater than the first threshold score.

1        60. (Amended) A method for processing insurance policies suspected of premium  
2        fraud, the method comprising:  
3        scoring each of a plurality of insurance policies with a predictive model to generate for  
4        each policy a score indicating a relative likelihood of premium fraud;  
5        determining for each scored policy an expected premium adjustment;  
6        ranking the scored policies according to their expected premium adjustments;  
7        selecting for a desk audit those policies having an expected premium adjustment  
8        exceeding a first threshold amount; and

9 selecting for a field audit those policies having a expected premium adjustment  
10 exceeding a second threshold amount, wherein the second threshold amount is  
11 greater than the first threshold amount.

1 61. (Amended) A method of developing a predictive model of insurance premium  
2 fraud, the method comprising:

3 collecting from at least one insurance company policy information for a plurality of  
4 insurance policies;

5 determining for each policy a scoring period for scoring the policy;

6 selecting a training set of policies;

7 deriving for each policy in the training set a plurality of variables from the policy

8 information and from other information relevant to policy premiums;

9 applying the derived variables to an untrained predictive model to train the predictive

10 model to produce a measure with respect to whether the policies are fraudulent or

11 non-fraudulent during their respective scoring periods ; and

12 selecting a subset of the derived variables for using in the predictive model, which

13 variables significantly contribute to a prediction of whether a policy is fraudulent

14 during its scoring period.

1 62. The method of claim 61, wherein the insurance policies are workers' compensation  
2 insurance policies, further comprising:

3 excluding from the training set policies for which no payroll is reported during the  
4 scoring period for the policy.

1 63. The method of claim 61, further comprising:

2 tagging each of the policies to indicate whether the policy is fraudulent, non-

3 fraudulent, or indeterminate; and

4 excluding from the training set policies which are tagged as indeterminate.